User Guide for High Resolution Climate Variable Downscaler
Version 1.0 (HRCD)

Contents

1. Use, disclaimer, and citation info

The code was developed by Alan Di Vittorio at the University of California, Berkeley (2008 - 2010). It is available for public, non-commercial use. This code has not been evaluated exhaustively and could contain bugs. By using this software you accept that you are using it at your own risk. Neither the author nor UC Berkeley are liable for anything you do with or to this code. It has been compiled and used successfully on Max OS X Darwin 9.# (currently 9.8) and Ubuntu 10.# (currently 10.10). Please use the following citation if you use this code (citation will be updated when it gets published):

Di Vittorio, A.V. and N.L. Miller (in review). Evaluating a modified point-based method to downscale cell-based climate data for high-resolution, gridded ecosystem modeling applications.

2. Requirements and installation

This software was written in C and was designed for unix/linux operating systems. In order to run this software you will need to download 3 data sets and install the NetCDF library (http://www.unidata.ucar.edu/software/netcdf/). Each data set is stored in a separate directory because each data set comprises several files. The file name formats are hardcoded in the software, with examples below, which shouldn't be problem for LSHRG and CGIAR data unless you change the file names when you download the data (or the names have changed on the available files). I renamed the GLOBE files so you will have to also. Each directory name containing the data files is also hardcoded in the software (also listed below), but these names are easy to change before you compile the code. In terrain.h, static char *glb_dir holds the GLOBE directory name and static char *srtm_dir holds the CGIAR directory name. In climate.h, static char *lshrg_path holds the LSHRG directory name. If you do need to change the file names in the code, most of the text portions of the names are also static chars in terrain.h and climate.h.

<u>The 3 required data sets are:</u>

NOAA NGDC GLOBE 1 km (30 arcsec) elevation data set:
http://www.ngdc.noaa.gov/mgg/topo/globe.html
The default storage directory is 'globe_dem_1km/'
My file name format is globe_r_c.dat, and these files are unzipped. The 'r' is the row number and the 'c' is the column number. The upper left tile is 'globe_0_0.dat' (corresponding to tile A) and there are 4 tile rows and 4 tile columns.
Hastings, D.A. and Dunbar, P.K., 1999. Global land one-kilometer base elevation (GLOBE) digital elevation model, documentation, volume 1.0. Key to Geophysical Records Documentation (KGRD) 34. National Oceanic and Atmospheric Administration National Geophysical Data Center, 325 Broadway, Boulder, Colorado 80303, U.S.A.

CGIAR–CSI v4.1 corrected SRTM 90 m (3 arcsec) elevation data set in Arc ascii format:
http://www.cgiar-csi.org/data/elevation/item/45-srtm-90m-digital-elevation-database-v41
The default storage directory is 'srtm_dem_90m/'
The file name format is srtm_cc_rr.zip where 'cc' is the column number and 'rr' is the row number. Columns go from 1-71, left to right, starting with -180° E as the left edge. Rows go from 1-24, top to bottom, starting with 60° N as the top edge. The code unzips and rezips files as needed. The upper left corner file name is 'srtm_00_00.zip.'
Jarvis, A., Reuter, H.I., Nelson, A. and Guevara, E., 2008. Hole-filled SRTM for the globe Version 4, CGIAR-CSI SRTM 90m Database. International Center for Tropical Agriculture, Cali, Columbia.

LSHRG daily 1x1 deg. 1948-2006 climate variable data set (including the elevation file)
http://hydrology.princeton.edu/home.php
The default storage directory is 'lshrg_clim_1deg/'
The file name format is variable_daily_yyyy-yyyy.nc where 'variable' is one of 9 variable abbreviations (see section 4 for variables, section 4.2 for abbreviations) and 'yyyy' is the year (same year for both yyyy). For example, the precipitation file for 2000 is named 'prcp_daily_2000-2000.nc. The elevation file is named elevation.nc.
Sheffield, J., Goteti, G. and Wood, E.F., 2006. Development of a 50-year high-resolution global dataset of meteorological forcings for land surface modeling. Journal of Climate, 19(13): 3088-3111.

To install the software:

First install the NetCDF 4 library
Put hrcd1.tar in your desired directory
Extract the source files using tar –xvf hrcd1.tar
This will create a directory called hrcd1 that should contain the following 21 files:
avg_clim.c
calc_coswt.c
calc_srad_hum.c
calc_srad_hum_iter.c
climate.c
climate.h
elev_adjust.c
get_past_clim.c
grids.c
grids.h
interp_clim.c
makefile
proc_clim.c
proc_terra.c
read_globe.c
read_srtm.c
terrain.c
terrain.h
user_guide_hrcd1.doc
wls_linreg.c
wls_linreg.h
You can rename or move the hrcd1 directory if you prefer
Copy netcdf.h into the hrcd1 directory
Set the USER and HOST variables in the makefile.
Set the path in the LDFLAGS_CLIM variable to point to your libnetcdf directory
If you are using a LINUX system you shouldn't have to change anything else
If you are using a UNIX system you need to comment the LINUX CC = gcc line and uncomment the UNIX CC = cc line. The LINUX CFLAGS variable should work for UNIX.
Within the source file directory type 'make all' at the command prompt.
The 'genterra' and 'genclim' binaries will be created within the source file directory.
You can change the destination of the binary files by adding a path in front of the $@ characters in the genterra and genclim makefile lines (the $@ characters denote the output binary file names, so you can change these also).
Don't worry about warnings stating that a variable is defined but not used. These warnings occur because not all included global variables are used in particular code files.

3. Overview

The High Resolution Climate Variable Downscaler (HRCD) takes a global, daily, coarse resolution (1x1°) climate variable data set and downscales the climate variables to a specified high-resolution grid (with daily time step). In the process a terrain grid (elevation, aspect, slope, east and west horizons) is generated that is organized in a row/column directory structure with each output cell having its own directory. The tile output directory contains the row/column directory structure. The tile directory can be the project directory itself or it can be one of several tile directories within the project directory. This structure allows a mosaic of tiles to be created in parallel within one project (for example, a global project computed as several tiles). The terrain data is output as a text file in each output cell directory can also be output as binary raster files in the tile output directory. The climate variable data can be output as ascii or binary files within each output cell directory or as NetCDF or day-specific raster files in the tile output directory.

HRCD comprises two C programs: genterra and genclim. Genterra needs to be run first to set up the output grid. Subsequently, genclim can be run several times to generate multiple sets of climate variables for a preexisting grid. The genclim input grid specification must match the genterra input grid specification of the preexisting grid.

4. Inputs and Outputs

<u>Inputs</u>

The inputs include a series of command line arguments in addition to the three required data sets listed above. These arguments are detailed below for each program (see Command line arguments) and are used to specify grid and processing parameters, input and output paths, and output file types. The two elevation data sets are used to generate the terrain grid. The GLOBE data set is for output grid cell sizes >= 1 km or 30 arcsec and the CGIAR data set is for output grid cell sizes < 1 km or 30 arcsec. The CGIAR data set covers only +-60° latitude, which is the geographic limit of output grids with cell sizes < 1 km or 30 arcsec. The LSHRG data set comprises an elevation file and 9 daily climate variables at the surface: downward long wave radiation, downward short wave radiation, precipitation, pressure, specific humidity, average temperature, maximum temperature, minimum temperature, and wind speed. There is one NetCDF file per year (1948-2006) for each variable.

<u>Outputs</u>

How to select output file options is detailed below for each program (see Command line arguments). HRCD calculates and outputs only entire years and only days of year 1-365. Here is a brief summary of the available output files:

4.1 Genterra Outputs

Tile directory

terrain_metadata.txt – always created
This file contains the grid specification and metadata for the binary raster files. It also contains the command line call to genterra and a comparison between the input and output values for the aggregated grid.

miss_terra_pix.dat – always created
Headerless raster file containing an output image of the number of missing data input cells per output cells. The file data starts at the upper left corner and is stored row by row with a 2-byte signed integer per output cell. This image should always contain all zero values, but I haven't checked the entire globe for both input data sets.

aspect.dat – 10*degrees
elevation.dat - meters
slope.dat – 10*degrees
east_hor.dat – 10*degrees; east horizon angle from horizontal
west_hor.dat – 10*degrees; west horizon angle from horizontal
These files are output if raster output is selected. The particular variable that are output depend on which ones were selected for general output. Each one is a headerless raster file containing an output image of the respective variable. The file data starts at the upper left corner and is stored row by row with a 2-byte signed integer per output cell.

Output cell directory

r#_c#_terrain_metadata.txt – always created; r=row, c=column
This file contains the latitude and longitude of the cell center and the area-weighted average aspect, elevation, slope, and the east and west horizons of the cell. Row and column indices begin with 0.

4.2 Genclim Outputs

Tile directory
All genclim tile directory output files are optional.

variable_yyyy_d.dat – created only for days listed in a file specified on the command line
Headerless raster file for named variable, year yyyy and day of year d (1-365). The file data starts at the upper left corner and is stored row by row with one single-precision floating point value per output cell (this is 4 bytes on a 32-bit processor). The geographic metadata are in terrain_metadata.txt in the tile (or project) directory. No rasters will be written if the specified file does not exist.

variable_daily_srfc_yyyy.nc – created for all years processed
NetCDF file for named variable and year yyyy. Metadata is included in the file.

All the variables are:

daylen	day length	s
dlwrf	downward long wave radiative forcing	W m^{-2}
dswrf	downward short wave radiative forcing	W m^{-2}
prcp	precipitation	cm
pres	average surface pressure	Pa
shum	average specific humidity	kg H_2O vapor kg^{-1} air
tas	average temperature	C
tday	average daytime temperature	C
tmax	maximum temperature	C
tmin	minimum temperature	C
vpd	average daytime vapor pressure deficit	Pa
wind	average wind speed	m s^{-1}

tilename_inputs.txt – outdated; only created when raster output is selected
You can ignore this file. It was used in development. It contains the input values used for some of the output values. It also contains some partially incorrect info about average values.

Output cell directory

styy_endy_clim_var_hin_intrp#X.txt – ascii file
styy_endy_clim_var_hin_intrp#X.dat – binary file
styy is the first year included in the file.
endy is the last year included in the file.
clim denotes climate file.
var is either 'all' or 'bgc,' depending on which set of outputs are selected.
intrp#X denotes how many points (#) are calculated and subsequently averaged within an output cell and what type of interpolation (X) is used for each calculated point. X = 'c' for modified cosine interpolation and X = 's' for gaussian interpolation. See genclim notes below for restrictions on these parameters.

You can select to output either an ascii file or a binary file. The binary file is much faster to write and access. Each file contains the number of days in the file, when the file was created, the latitude and longitude of the output cell center, the elevation of the output cell, and the variables and their units. Each daily record includes the year and day of year.

The binary file also contains info on the number of header bytes (512), the record size (one record per day), the total number of data bytes, and the data types for the variables within the file. The binary file format begins with 512 header bytes that are ascii characters containing the information described above. The remaining data are in day-by-day order. The first two fields of each daily record are short integers (year and day of year) and the remaining fields are single-precision floating point values. These values are usually 2 bytes and 4 bytes, respectively, on 32-bit processors.

5. Genterra

5.1 Important use notes

This code outputs values that correspond to a cell area, rather than a specific point

Set up grids so that cell and tile boundaries correspond to following global limits, if grids are near these edges:
    +-90 lat and +- 180 lon deg
    sin proj x range is -21600000 to 21600000 m; y range is -10800000 to 10800000 m

Only the inputs and a base directory need to exist
    The output grid path and directory structure will be created if it doesn't already exist

The largest output tile this program will accomodate contains 6,480,000 pixels
    this is equivalent to a global geographic grid with .1 degree (6') cell size
    this is about 25000x25000 cells with 1 km cell size (covers ~20x20 sinusoidal projected degrees, or 2x2 modis tiles)

A count of missing input cell values for each output cell is output as a binary file in the same format as the terrain binary files

Currently the output cell size is equal projected meters or arcseconds in both dimensions

There is a minimum cell size of 100 meters for sinusoidal output, and 3 arc-seconds for geographic output
    This is based on having a minimum of 1 input point to average

There is a maximum cell size of 1200 km for sinusoidal output, and 10 degrees for geographic output
    This is to restrict data structure sizes for processing (it could be larger, but we don't need larger)

Notes about sinusoidal projection:
    Currently the lon origin can be set, but a tile cannot cross the +-180 line
    Inverse projection is true to geographic coordinates, but projected area and scale are not equal to ground area and scale
    because the projection assumes a spherical earth
    Thus the scale varies within the projection, and nowhere is it exact with ground measurements
    Projected cells are bounded by integral arc-seconds of projected dd
    Projected scale is: 100m = 3 arc-seconds at the equator and along meridians

Thus 100 projected meters is 92.66254331 spherical surface meters (using AVE_ER = 6371007.181)

I have read that the difference between spherical and spheroidal distance can be up to 0.55% at the equator, and in general averages to about 0.3%

My estimated max error for area is 0.45% at poles and 0.22% at equator; with respect to AVE_ER.

However, the error between spherical areas usig WGS-84 max and min radii is 0.67% with respect to AVE_ER.

These estimates are based on spherical areas, assuming that the spheroidal curvatures differ negligably from the spherical.

This means that very large cell sizes might have higher error than 0.45%.

Note about geographic coordinates:

Do not calculate input arc-second values from decimal degrees: 0.666 is not the same as 2400 arc-seconds

The output values are within-out-cell-averages of the values calculated at each input point

Values at the output cell boundaries are included in the average

For output cell sizes less than 30" or 1 proj km, the latitude range is restricted to +-60 degrees

This is the limit of SRTM data, and I haven't implemented any interpolation of GLOBE data

Missing elev values at the +60 lat edge of SRTM are set equal to immediate southern neighbors

Missing elev values at the -60 lat edge of SRTM are set equal to GRID_OCEAN

There should be only two input rows at each lat edge with missing elev values: +60, -60, 60+3sec, -60-3sec

Average slope and aspect are calculated from the average gradient vector

Aspect is the opposite direction of the gradient vector

Slope as angle is the arctan of the magnitude of the gradient vector

A gradient vector with 0 slope will return an aspect = 90 = E

For east and west horizon, the angle values are averaged

Output cells are set to ocean (-5555) if input ocean area >= a fraction of non-missing out-cell area

OCEAN_FRAC is currently set to 1.0,

which means that a single input land cell will set an output cell to land

Output cells are set to nodata (-9999) if the missing data input area >= a fraction of total out-cell area

MISS_FRAC is currently set to 0.5,
which means that if 50% or more inputs cells have no data then the output cell will be set to nodata

The default compiled name of this code is “genterra"

5.2 Command line arguments

The genterra command line arguments are organized as follows (9 or 10 required arguments (depending on projection) and 1 optional):

genterra -[g<cell_size> <ul_lat> <ul_long> <lr_lat> <lr_long> 'OR' a<cell_size_m> <lon_origin> <ul_y> <ul_x> <nrows> <ncols>] -[r 'OR' p 'OR' b] -esadu <source_data_path> <output_path> <tile_name (opt)>

-g outputs a geographic projection delimited by the input upper left and lower right corners
cell size and lat long values are in decimal arc-seconds
lat range is -324000 to 324000 (-90 to 90 degrees); lon range is -648000 to 648000 (-180 to 180 degrees)
these values define the actual corners of the corner cells

-a outputs a sinusoidal equal area projection with nrows*ncols cells of size=cell_size_m the ul corner input is ul image corner in sinusoidal projected meters (see MODIS land tiles)
x range is -21600000 to 21600000; y range is -10800000 to 10800000
cell_size_m is in sinusoidal projected meters
lon_origin is in decimal arc-seconds and is the geographic lon origin of the sinusoidal projection (range: +-648000)
projected meters and projected dd are 0,0 at lon_origin
lat origin is always the equator, with projected meters = 0 and projected dd = 0

-r produces binary rasters of the outputs, with a text metadata file
this was supposed to be controlled by –esadu, but it doesn’t work correctly
these rasters are written row by row starting at the ul corner
the data are 16-bit signed integers

-p produces terrain text files in each point directory of the project bgc data structure, with a tile text metadta file as in -r
the lat and long values are at the cell center
elevation is in meters and slope, aspect, east and west horizons are in degrees

-b produces both rasters (with metadata) and terrain text files
this raster output was supposed to be controlled by -esadu below

-esadu: you must enter this argument with all 5 letters for the code to run, regardless of

whether you use –p, -r, or -b
This was intended to select any combination of five variables for raster output,
But it is not working correctly
e=elevation in meters
s=slope in 10*degrees
a=aspect in 10*degrees; 0=N, 90=E, 180=S, 270=W
d=west horizon in 10*degrees
u=east horizon in 10*degrees

source_data_path is the directory containing the directories of source elevation data

output_path is the directory containing the tile directories (or row/col dirs, see tile_name)
this is the project path
This path is created if it doesn't exist

tile_name is the optional name of the tile directory that contains the row(lat) and column (lon) directories for the project
this is the directory where the raster files will be written
this directory contains one 'tile' of the globe
do not include '/' at the end of tile_name
if no tile name is given then the row and column directories are put in output_path
the tile directory does not have to exist -- it will be created if it is needed

5.3 Example

genterra -g150 138825 -433875 138525 -433575 -b -esadu /home/hrcduser/geodata/ /home/hrcduser/projects/example swca_150sec

This call to genterra creates a 2x2 cell terrain grid with geographic projection and 150 arcsec cell size. The upper left corner of the grid is at 138825” lat and -433875” lon and the lower right corner of the grid is at 138525” lat and -433575” lon. Raster images of the variables are output in the tile directory in addition to the text files output within each output cell directory (-b). All five letters ‘esadu’ have to be present in this argument for the code to run. The elevation data set directories (one for GLOBE and one for CGIAR data) are in source directory ‘/home/hrcduser/geodata/’ and the tile ‘swca_150sec’ will be a directory in ‘/home/hrcduser/projects/example’ and will contain the row/column directory structure. Rows and columns begin with index 0.

6. Genclim

6.1 Important use notes

Only full years are calculated with day of year 1-365
day 366 in leap years is not used or calculated
Input data needs to be in full year increments

IMPORTANT: The directory structure referred to in the arguments must already exist, and each cell (column) directory must contain a text file with cell terrain information. This text file must be output by genterra, or created with the exact same format and content.
This text file is named: r#c#_terrain_metadata.txt; where # is replaced by row or col index, starting with index 0

This code outputs values that correspond to a cell area, rather than a specific point

The largest output tile this program will accomodate contains 6,480,000 pixels
  this is equivalent to a global geographic grid with .1 degree (6') cell size
  this is about 25000x25000 cells with 1 km cell size (covers ~20x20 sinusoidal projected degrees, or 2x2 modis tiles)

Currently the output cell size is equal projected meters or arcseconds in both dimensions

There is a minimum output cell size of 100 meters for sinusoidal output, and 3 arc-seconds for geographic output
  This is based on having a minimum of 1 input point to average terrain data (see terrain.c)

There is a maximum output cell size of 1200 km for sinusoidal output, and 10 degrees for geographic output
  This is to restrict data structure sizes for processing (it can be larger, but we don't need larger)
  Climate data is interpolated up to output cell size of (not including) max input cell dimension, and averaged for >= this value

Note about sinusoidal projection:
  Inverse projection is true to geographic coordinates, but projected area and scale are not equal to ground area and scale
  because the projection assumes a spherical earth
  Thus the scale varies within the projection, and nowhere is it exact with ground measurements
  Projected cells are bounded by integral arc-seconds of projected dd
  Projected scale is: 100m = 3 arc-seconds at the equator and along meridians
  Thus 100 projected meters is 92.66254331 spherical surface meters (using AVE_ER = 6371007.181)

I have read that the difference between spherical and spheroidal distance can be up to 0.55% at the equator, and in general averages to about 0.3%

  My estimated max error for area is .45% at poles and .22% at equator; with respect to AVE_ER.
  However, the error between spherical areas usig WGS-84 max and min radii is .

67% with respect to AVE_ER.
These estimates are based on spherical areas, assuming that the spheroidal curvatures differ negligably from the spherical.
This means that very large cell sizes might have higher error than .45%.

Note about geographic coordinates:
Do not calculate input arc-second values from decimal degrees: 0.666 is not the same as 2400 arc-seconds

The default compiled name of this code is “genclim”

6.2 Command line arguments

The command line arguments are organized as follows (17 or 18 required arguments (depending on projection), 1 optional argument):

genclim -[g<cell_size_seconds> <ul_lat> <ul_long> <lr_lat> <lr_long> 'OR' a<cell_size_m> <lon_origin> <ul_y> <ul_x> <nrows> <ncols>]
-[r 'OR' nc 'OR' pa 'OR' pb 'OR' ba 'OR' bb 'OR' bnca 'OR' bncb] -[i 'OR' m]
-[f 'OR' d<prcp_sm_box> 'OR' w<prcp_sm_box>] -[c<num_interp> 'OR' s<num_interp>]
-[l 'OR' n] <st_yr> <st_day> <end_yr> <end_day> <out_day_file> <source_data_path>
<output_path> <tile_name (opt)>

-g outputs a geographic projection delimited by the specified upper left and lower right corners
this needs to match the genterra call that created the grid
cell size and lat long values are in decimal arc-seconds
lat range is -324000 to 324000 (-90 to 90); lon range is -648000 to 648000 (-180 to 180)
these values define the actual corners of the corner cells

-a outputs a sinusoidal equal area projection with nrows*ncols cells of size=cell_size_m
this needs to match the genterra call that created the grid
the ul corner input is ul image corner in sinusoidal projected meters (see MODIS land tiles)
x range is -21600000 to 21600000; y range is -10800000 to 10800000
cell_size_m is in sinusoidal projected meters
lon_origin is in decimal arc-seconds and is the geographic lon origin of the sinusoidal projection (range: -180 to +180)
projected meters and projected dd are 0,0 at lon_origin

-r produces binary rasters of the climate outputs designated by -[i 'OR' m], with an outdated diagnostic file; in project or tile directory
these rasters are written row by row starting at the ul corner
the data are single-precision floating point values (32 bits for 32-bit processors)
the geographic metadata are in terrain_metadata.txt in the tile (or project) directory

a text file named <out_day_file> must be in the tile directory containing days for which to create binary climate rasters, see below for format; if this file is absent, no binaries will be written

-nc produces NetCDF files of the climate outputs designated by -[i 'OR' m]; in project or tile directory

Each variable has one file per year; each variable has 3 dimensions: lon, lat, time
the data are single-precision floating point values (32 bits for 32-bit processors)

-pa produces climate text files in each point directory of the project bgc data structure

the output variables are designated by -[i 'OR' m]
precision is 2 decimal points (whole second for daylength)
There will be four header lines:

line 1: number of days (records) in the file, and date and time of file creation
line 2: lat, lon, and elevation
line 3: variable names as below
line 4: variable units as below

Followed by one daily record per line
The bgc output (-m) has the following 9 columns:
year, day, tmax (C), tmin (C), tday (C), precip (cm), vpd (Pa), srad (W/m^2), daylen (seconds)
The lshrg output (-i) can have up to the following 14 columns, if they are all available:
year, day, tmax (C), tmin (C), tave (C), precip (cm), shum (kg wv/kg air), pres (Pa), dswrf (W/m^2), dlwrf (W/m^2), windsp (m/s), tday (C), vpd (Pa), daylen (seconds)

-pb produces climate binary files in place of the text files described for -pa

There is a 512 byte ascii header with the following info:

# of header bytes, # of days in file, time of file creation, lat, lon, elevation, variable names and units as above, and the byte sizes of the variables

The data is written day-by-day after the header, with each day having the variables in the order listed above
year and day are type short, and the rest are type float

-ba produces both rasters and point text files

raster output controlled by <out_day_file> and -[i 'OR' m]

-bb produces both rasters and point binary files

raster output controlled by <out_day_file> and -[i 'OR' m]

-bnca produces both netCDF files and point text files

-bncb produces both netCDF files and point binary files

-i output all avalable variables:

tmax (C), tmin (C), tave (C), precip (cm), shum (kg wv/kg air), pres (Pa), dswrf (W/m^2), dlwrf (W/m^2), windsp (m/s), tday (C), vpd (Pa), daylen (seconds)

-m output the bgc model suite of variables:
tmax (C), tmin (C), tday (C), precip (cm), vpd (Pa), srad (W/m^2), daylen (seconds)

-f uses modified cosine interpolation between four immediate neighbor input points; non-elev-adjusted (flat)
this is not recommended because it produces poor results

-d uses mtclim with the daymet algorithm to calculate elev-adjusted climate for interpolation and averaging; (dry) no humidity input, tdew=tmin
this is not recommended because the results are not as accurate as when using the humidity input
<prcp_sm_box> is 0 or an odd integer value >= 5 and <= 9
this is recommended to be 0 because precip smoothing reduces accuracy for the LSHRG data set
this sets a two-sided linear ramp temporal smoothing filter for precipitation values; max weight at center, zero at the edges
so it has to be >= 5 for any smoothing to occur

-w uses mtclim with the daymet algorithm to calculate elev-adjusted climate for interpolation and averaging; (wet) with humidity input
if humidity inputs are not available for a particular year, mtclim will automatically use the non-humidity-input algorithm for this year
this is the recommended setting (-w0)
<prcp_sm_box> is 0 or an odd integer value >= 5 and <= 9
this is recommended to be 0 because precip smoothing reduces accuracy for the LSHRG data set
this sets a two-sided linear ramp temporal smoothing filter for precipitation values; max weight at center, zero at the edges
so it has to be >= 5 for any smoothing to occur

-c use modified cosine algorithm for spatial interpolation of elev-adjusted values
this is the recommended setting (-c1) because the gaussian interpolation is poor, buggy, or both
<num_interp> is number of interp rows within output cell; the number of interp cols currently equals the number of interp rows; range= 1 to 9
!!-use only num_interp = 1 because additional points do not add anything because they are at same elevation!!

-s use daymet's gaussian interpolation for for spatial interpolation of elev-adjusted values
this is not recommended because the results are poor, buggy, or both
<num_interp> is number of interp rows within output cell; the number of interp cols currently equals the number of interp rows; range= 1 to 9

!!-this doesn't seem to be working correctly!!
!!-use only num_interp = 1 because additional points do not add anything because they are at same elevation!!

-l uses 1deg climate data from LSHRG; data type for these netcdf files is float
this is your only option

-n uses NASA GMAO DAS data -- not implemented

st_yr is the year (yyyy) to start processing climate data

st_day is the day of the year (1-366) to start processing climate data
not implemented: currently hardcoded to 1, but still need the argument

end_yr is the year (yyyy) to end processing climate data

end_day is the day of the year (1-366) to end processing climate data
not implemented: currently hardcoded to 365, but still need the argument

out_day_file is the name of the text file that contains the days for which to output raster climate data
each line contains two values separated by whitespace: yyyy day(1-365); year and day
this command line argument must be present, even if it is not used -- can insert a dummy character

source_data_path is the directory containing the directories of source climate data

output_path is the directory containing the tile directories (or row/col dirs, see tile_name), also known as the project path
this needs to have been created by genterra, so this has to match the genterra call
this is tile directory if no tile name is given

tile_name is the optional name of the tile directory that contains the row(lat) / column (lon) directories
this needs to have been created by genterra, so this has to match the genterra call
this is the directory where the raster files will be written
this directory contains one 'tile' of the globe
do not include '/' at the end of tile_name
if no tile name is given then it is assumed that the row and column directories are in the output_path
row directories are in tile (or project) directory, and col directories are in the row directories

6.3 Example

genclim -g150 138825 -433875 138525 -433575 -bnca -i -w0 -c1 -l 2001 1 2003 365 nofile /home/hrcduser/geodata/ /home/hrcduser/projects/example swca_150sec

This genclim call assumes that the example genterra call in section 5.3 has already been successfully completed so that the specified grid directory structure already exists. An ascii file will be written to each cell directory and a NetCDF file will be written to the tile directory '/home/hrcduser/projects/example swca_150sec' (-bnca). All available climate variables will be output (i) and calculated using humidity input with no precipitation smoothing (-w0) and cosine interpolation with one point per output cell (-c1). The source climate data are LSHRG (-l) and the years to calculate are 2001 – 2003. The argument 'nofile' is a placeholder for the file name associated with raster image outputs. The source directory '/home/hrcduser/geodata/' contains the LSHRG data directory and the project and tile directories math those in the genterra example call.